AD_____

Award Number: DAMD17-02-1-0499

TITLE: Cloning and Characterization of Expanded CAG-Repeat Containing Sequence(s): Identification of Candidate Breast Cancer Predisposition Gene(s)

PRINCIPAL INVESTIGATOR: Hilmi Ozcelik, Ph.D.

CONTRACTING ORGANIZATION: Mount Sinai Hospital
Toronto, Ontario, M5G 1X5 Canada

REPORT DATE: June 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2004	Annual (1 May 2003 - 30 Apr 2004)

4. TITLE AND SUBTITLE
Cloning and Characterization of Expanded CAG-Repeat Containing Sequence(s): Identification of Candidate Breast Cancer Predisposition Gene(s)

5. FUNDING NUMBERS
DAMD17-02-1-0499

6. AUTHOR(S)

Hilmi Ozcelik, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Mount Sinai Hospital
Toronto, Ontario, M5G 1X5 Canada

E-Mail: ozcelik@mshri.on.ca

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Original contains color plates: ALL DTIC reproductions will be in black and white

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Genetic factors have been shown to influence the risk of developing breast cancer. To date only two breast cancer predisposition genes, BRCA1 and BRCA2, have been identified. A wide spectrum of BRCA mutations has been found in breast cancer families. However, in the majority of breast cancer cases BRCA1 or BRCA2 mutations were not detected, suggesting the presence of other breast cancer predisposition genes. Trinucleotide repeats are present widely in human genome. Expansions of some trinucleotide repeats have been recognized to be the cause of several genetic disorders. In a previous study using the Repeat Expanded Detection (RED) method, we detected expanded (CAG) repeat of repeat sizes of over 144 in 2.4% of breast cancer cases in a sample of 212 patients. No expansion of this magnitude has been detected in 196 population controls samples. Since RED does not provide information about the location of expanded repeats, it is necessary to clone and localize them. In this stage of the project we have developed and optimized an efficient cloning and screening strategy using a combination of dynabead enrichment microsatellite isolation protocol and the RepeatArray system. We also validated the results by sequencing the inserts. We have shown that applying the dynabead enrichment for a second round dramatically increases the percentage of CAG containing clones (around 90%). We have also detected a large GAG/CTG containing fragment of 65 repeats suggesting that our system is able to identify large repeat containing inserts. We have also established RepeatArray technology, which provides the distinction between small and large repeat-containing fragments and provides a high throughput pre-screening tool and eliminates the need for excess sequencing. Currently, we are at the stage of further validation and application of the strategy to breast cancer.

14. SUBJECT TERMS
Breast cancer, CAG/CTG repeats, Rapid Expansion Detection, Trinucleotide Repeat Expansions

15. NUMBER OF PAGES
17

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Cover.....1

SF 298.....2

Table of Contents.....3

Introduction 4

Statement of Work.....5

Body.....7

Key Research Accomplishments.....14

Reportable Outcomes.....15

References.....17

I. INTRODUCTION:

Expression of an inherited disease gene does not always follow a Mendelian inheritance pattern. Since the turn of the century, it has been observed that the transmission of certain disease genes from one generation to the next is associated with an increase in the severity of the disease symptoms and/or a decrease in the age of onset. This clinical phenomenon is called anticipation and it was observed with many inherited diseases, especially ones involving neurological disorders. However, the absence of a satisfactory physical explanation for anticipation at the molecular level resulted in much controversy in the past over the acceptance of the concept of genetic anticipation.

Advances from molecular genetic analysis of many of the inherited human diseases in which anticipation has been observed has recently attributed anticipation in these diseases to a new class of dynamic mutations, characterized by trinucleotide repeat expansions in the locus where the disease genes have been mapped. The effect depends on the location of the repeats relative to the gene and the type of repeats (for review see Sanjeeva et. al., 1997, Margolis et al., 1999, Vincent et al., 2000). The molecular consequences that result from trinucleotide repeat expansions and the mechanism by which they lead to pathology may be quite diverse. In general however, trinucleotide repeat expansions have been shown to perturb either the structure and function (type I mutations) or the expression (type II mutations) of the affected gene (for review see Sanjeeva et. al., 1997, Margolis et al., 1999, Vincent et al., 2000). For example, expansion of a (CAG)n-repeat in the coding region of the Huntington's disease gene allows expression of an altered protein which contains an expanded polyglutamine region, resulting in altered conformation, processing and general physical properties of the protein function of the product (Trottier et al., 1995). Alternatively, expansion which results in 200-2000 CGG repeats in the 5' untranslated region of the Fragile X syndrome (FRAXA) gene result in the loss of expression of FRAXA mRNA (Pieretti et. al., 1991). As the repeat expands with transmission to the next generation, the CGG repeats become more methylated reducing transcription of the FRAXA gene. The nature of this dynamic group of mutations, which can involve very large amplification of trinucleotide repeats, renders the sequence unstable during meiosis. This results in intergenerational instability of the length of the trinucleotide repeat. It is not known why repeats which exceed a critical value are unstably transmitted to succeeding generations with a tendency towards expansion of trinucleotide repeats. There is however, a very clear association between longer expansions at the disease locus in the succeeding generations and earlier clinical manifestation (Sanjeeva et. al., 1997, Margolis et al., 1999, Vincent et al., 2000). Anticipation, therefore, is now commonly accepted as a hallmark of the inheritance of an amplified trinucleotide repeat expansion mutation. Thus far, at least 12 genetic diseases (mostly muscular or neurological disorders) have been attributed to expansions of trinucleotide repeats in the loci containing the disease gene. These diseases are characterized as having increasing copy numbers of the unstable expanded sequences with subsequent generations.

One can envision that expansion of trinucleotide repeats are not restricted to neuro-muscular disorders, and they probably represent a novel class of dynamic mutations causing various human diseases. A study by O'Donovan et al., 1996,

provided an intriguing finding suggesting that repeat expansions have a wide spread role in common human diseases. This study has shown that older, healthy individuals have generally shorter CAG-repeat lengths than their younger counterparts. The demonstration of decreased, genome wide repeat-copy number with age in healthy populations suggests that the dynamic mutation could have a wide-spread role in human susceptibility to disease. These findings also suggest that trinucleotide repeat diseases do not display a single major gene inheritance.

II. Statement of Work

Technical Objective 1: Cloning of Gene Sequences with CAG-Repeat Expansion.
(Four parallel cloning procedures will be performed simultaneously using 4 pools each including DNA mixed from 2 RED positive cases)

Task 1: Months 1-4	Enrichment by digestion of DNA samples and separation on agarose gels. Slicing (2-4mm), and extracting the DNA from each slice (approximately 200-400 slices). RED analysis of all 200-400 slices to locate the fragments with expansion.
Task 2: Months 4-6	Cloning into ZapII system, electroporation into E.Coli, and amplification. Secondary enrichment by pooling: Plating the bacteria, pooling and amplification of pools. Extraction of DNA (~100-200) and RED analysis (~ 100-200).
Task 3: Months 6-8	Transformation of the RED positive DNA into E.Coli CJ236 strain and generation of ssDNA. Isolation of the ssDNA and production of dsDNA containing only CAG inserts CAG-probes and the primer extension method. Electroporation into E.Coli, amplification, and plating. Selection and amplification of individual clones. Extraction of DNA from individual clones (~100-200).
Task 4: Months 8-10	RED analysis on the DNA extracted from the individual clones (~100-200). Sequencing of inserts from RED positive clones. Identification of clones with large repeat sequences.
Task 5: Months 10-11	Designing PCR primers for every sequence with large CAG-repeats identified through cloning. Screening of all the cases originally detected to have the CAG-repeat expansion (detected by RED) by PCR and sequencing.

Technical Objective 2: Identification and Characterization of Gene(s) containing or Flanking Expanded CAG-Repeats

Task 1: Months 11-12	Designing PCR protocols for all the sequences identified to have large CAG-repeats through cloning and optimization of their detection through microsatellite analysis. Sequencing of a small panel of control specimens to identify repeats with

varying sizes to be used as size controls on microsatellite gels.

Task 2: Months 12-17 — Microsatellite analysis of 100 control specimens for every repeat cloned to determine the allelic frequency of each repeat. It is expected than several regions containing large CAG-repeats will be identified from cloning of RED positive DNA from 8 breast cancer cases.

Task 3: Months 17-22 — Extraction of sequence data using different sequence database sources including GenBank, EST and STS databases and putting together the pieces of information to find longer sequence information. Mapping of the CAG repeats using public genetic map information, and identify other genes located around the CAG-repeat. Locating the exact position of CAG repeat in relation to the structure of the genes identified.

Task 4: Months 20-24 — Obtaining the sequence data from the databases and interpretation of this information with the presence of an extensive literature search. If necessary, wet lab work will be carried out to clarify the position of CAG-repeat in relation to the structure of gene(s) identified. Search for information published on the loci/genes found and their association with cancer specifically with breast cancer. Writing of manuscripts to be proposed to peer reviewed journals.

III. Body

A. Overview

Our goal is to clone and characterize expanded CAG repeat containing sequences in breast cancer patients. These genes may represent breast cancer predisposition genes. Most of the reported methods utilized for library construction and cloning of the large repeat containing fragments require the use of large amount of genomic DNA as a starting template (Yuan et al, 2001; Vincent et al, 2000; Koob et al, 1998). Since patient material is a limited source, we were specifically interested in developing an efficient cloning strategy that enriches for long CAG repeat containing fragment with the minimum use of starting genomic DNA. Long trinucleotide repeats tends to form special secondary structures, which reduce the efficiency of cloning of DNA fragments containing expanded repeats. Therefore, cloning of long repeat containing fragments is cumbersome. (Koob et al, 1998, Sanpei et al, 1996)

As described in the previous annual report, the isolation procedure of trinucleotide repeats from genomic DNA has been established and the efficiency of the method has been evaluated. The protocol, which was modified from (Inoue et al, 1999) is based on the DNA polymerase reaction, which is restricted with only three nucleotide substrates and primed by a biotinylated probe. Sequences containing trinucleotide repeats are then isolated by a streptavidine biotin trapping method. Using this method, we had successfully detected the repeat containing fragments, however at a low efficiency (60%). More recently, we have also applied the Microsatellite Isolation with Dynabeads protocol which was developed by Travis Gelnn at the Department of Biological Sciences in the University of South Carolina who has refined the approach described in Hamilton et al. (1999), which is a similar, modified version of the previously used protocol. This procedure was found to be more efficient for the purpose of our project (see below). We also developed the RepeatArray technique as an efficient and high throughput tool for the detection of long CAG repeat containing clones.

B. Progress

1. The recent Microsatellite Isolation with Dynabeads Cloning Method

We have established this method, which consists of eleven main steps **(Figure 1).**

a) Restriction Enzyme Digestion

During this step the DNA is fragmented into fragments of approximately 500 base pair (500 bp) fragments. These fragments are then inserted into a plasmid. Fragments of this size are small enough to sequence easily, yet still retain a high probability of having enough DNA flanking the microsatellites that primers can be designed. The restriction enzymes digestion is carried by using RsaI and BstU I.

b) Ligating Linkers to DNA Fragments

In this step double stranded linker is ligated to both ends of each DNA fragment. The linkers provide the primer binding sites for subsequent PCR steps. They also provide sites to ease cloning of the fragments into the vectors to be used subsequently. The linkers are, therefore, compatible with the restriction sites in the vector's multiple cloning

sites. The linker also incorporates a GTTT "pig-tail" to facilitate non-template A addition by Taq DNA polymerase during PCR which can be used for TA cloning

c) PCR Amplification of DNA Fragments

During this step the DNA fragments which has been ligated to the linkers are amplified in a regular polymerase chain reaction using one of the ligated linkers as a primer in order to increase the copy number of digested genomic DNA fragments prior to enrichment for CAG repeats.

d) Dynabead Enrichment for Microsatellite-containing DNA Fragments

The goal of this step is to capture DNA fragments with CTG repeat sequences complementary to the CAG Probe and wash away all other DNA fragments. A biotinylated $(CAG)_{10}$ probe is annealed to the amplified DNA fragments produced in the previous step. Then a primer extension reaction is performed using only three dNTPs (dATP, dCTP and dGTP). The resulting primer extension reaction is then hybridized to dynabeads and captured using a magnetic particle collecting unit in order to enrich for CAG repeat containing fragments and the other fragments are washed away. Finally the CAG enriched fragments are released from the dynabeads in ssDNA form to be used in the next step.

e) PCR Recovery of Enriched DNA

The purpose of this step is to increase the amount of CAG enriched DNA by performing a PCR reaction using the linker as a primer.

f) Further Enrichment for Microsatellite-containing DNA Fragments

In this step the DNA fragments are further enriched for CAG repeat containing fragments by repeating steps (d) and (e) using PCR products from step e as a template.

g) Ligating Enriched DNA into Plasmids

The aim of this step is to incorporate the enriched/recovered DNA (amplified in step e) into a cloning vector. We have used the Invitrogen TOPO TA Cloning® Kit containing pCR®2.1-TOPO® with TOP 10 cells (catalog K4500-40) and we followed the directions supplied with Invitrogen's Topo-TA cloning kit exactly.

h) Transforming Plasmid DNA

The purpose of this step is to incorporate the enriched/recovered DNA (amplified in step e) and the cloning vector into a bacterial host. The plasmids from step (e) were transformed into TOPO 10 competent cells provided with TOPO TA Cloning Kit (Invitrogen) following the manufacturer instructions.

i) Library construction

Transformation products from step (h) are spread onto petry dishes and incubated overnight at $37^{O}C$.

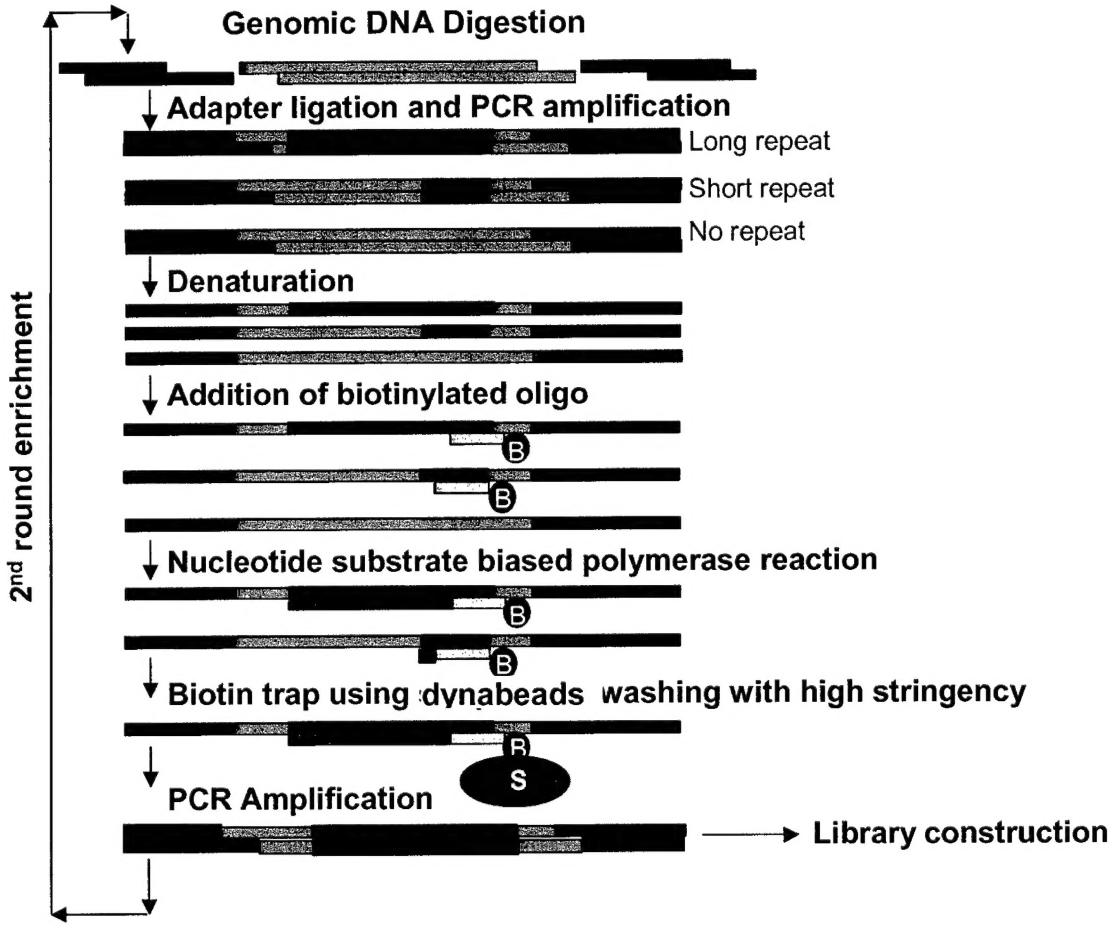
j) Colony picking and PCR amplification

White colonies are inoculated into 100 µl of LB/Amp medium and incubated for 6-8 hours. Then PCR is performed using the M13 forward and reverse primer set (from the vector sequence) and using 1 µl of the LB culture as a DNA template without the need of performing any plasmid DNA isolation procedure.

k) Screening for long CAG repeat containing inserts

PCR products from step (j) were used for performing DNA sequencing and/or RepeatArray screening in order to find long CAG repeat containing inserts as described below.

Figure 1. Schematic representation of enrichment for long CAG repeat containing fragments by using Dynabeads



2. Determining the Efficiency of the Cloning Protocol Using Sequencing

In order to test the efficiency of our protocol in identifying CAG containing repeats, we have randomly selected 80 clones for DNA sequencing. 48 of them were subject to one round of dynabead enrichment (single enrichment) and the remaining 32 samples were subject to two rounds of dynabead enrichment (double enrichment). Automated sequencing was performed using the M13 forward primer and the sequences were analyzed and the number of inserts containing (CAG) repeats was determined **(Table 1 and Figure 2).** Among the 48 single enrichment samples only 20 samples (42%) contained CAG repeats ranging in size from 3 to 46 repeats. On the other hand, 28 samples out of the 32 double enrichment samples contained CAG repeats ranging in size from 4 to 65 units. This show that the second enrichment step has drastically improved the detection of fragments with CAG repeats. The identification of the $(CAG)_{65}$ containing fragment suggests that our system is able to identify large repeat containing inserts. Compared to our previous method this method found to be more efficient (60% VS 88%).

Table 1. Percentage of (CAG) containing inserts after single and double dynabead enrichment

Clones	Single Enrichment	Double Enrichment
Number of Inserts	48	32
Inserts with (CAG)	20	28
Percentage of Inserts with (CAG)	42%	88%
Minimum repeat length	(CAG)3	(CAG)4
Maximum repeat length	(CAG)46	(CAG)65

3. RepeatArray Technology

In the previous section we showed that using microsatellite enrichment by dynabead and DNA sequencing, we were able to increase the percentage of repeat containing inserts up to 88%. Since we are interested in relatively longer CAG repeats, we developed an efficient pre-screening technology to discriminate long from short and non-repeat containing sequences. This method drastically reduces the effort for sequencing, and increases the efficiency of detecting large repeats. We have established this DNA array system, which is based on stabilizing the PCR products from step (J) in the previous section onto a slide and hybridizing it with fluorescently labeled (CAG) probe then read it by an array scanner **(Figure 2).**

a) Spotting and Post-Spotting Processing

PCR products (5µl) were mixed with 5 µl of Micro Spotting Solution Plus in a 348 wells plate and are given to the microarray facility to be printed on slides. Then the slides are processed and cross-linked using the standard array post-spotting processing protocol and the slides were kept in a humid chamber for further analysis.

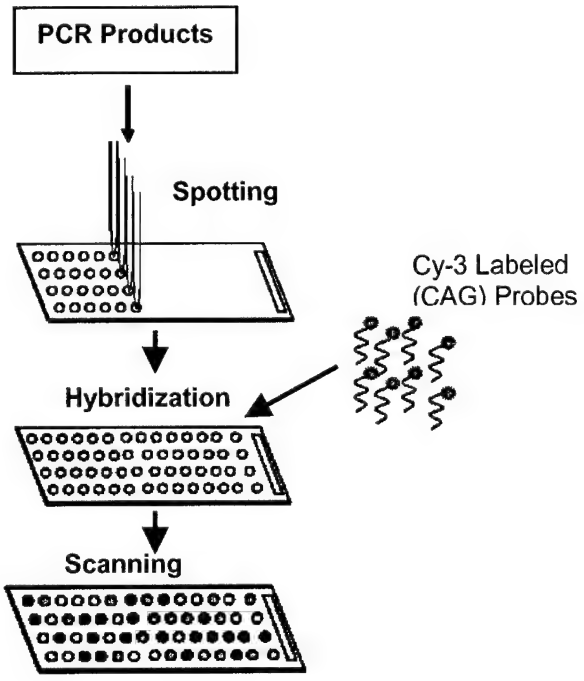
b) Hybridization with (CAG) Probes

We used two cy-3 5' labeled probes; $(CAG)_7$ and $(CAG)_{17}$. In order to obtain the optimal probe concentration and washing conditions we performed an initial hybridization and washing experiment using a range of probe concentration and different washing treatment until we obtained the optimal conditions.

c) Scanning and Analyses

Array slides were read using the GenePix 4000 B scanner with wavelength of 532 nm. The scanned images were analyzed using GenePixPro software (Axon Laboratory) following the manufacturer instructions. The mean intensity of array spots was proportional to the (CAG) length in the PCR product for both $(CAG)_7$ and $(CAG)_{17}$ probes **(Figures 3 and 4).** As shown in the figures the dynabead double enrichment protocol combined with RepeatArray procedure have been optimized and ready to use as a rapid, efficient and reliable system for localizing long (CAG) repeats in patient samples.

Figure 2: Schematic representation of the RepeatArray protocol.



In order to test the usefulness of the method, we have spotted a panel of PCR products containing a range of repeat sizes ranging from $(CAG)_0$ to $(CAG)_{45}$. In this experiment, we have shown that the signal intensity of both large and the small probe is relative to the size of the repeat content **(Figure 4)**. Variation from this trend is expected

if the quality of the PCR product is low. We are currently working on normalizing for this variable.

Figure 3: Arrays spots intensity of variable length CAG containing PCR products hybridized with (A) $(CAG)_7$ and (B) $(CAG)_{17}$ probe.

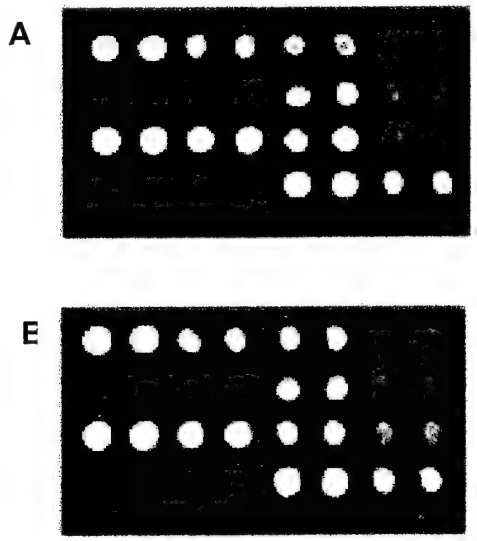
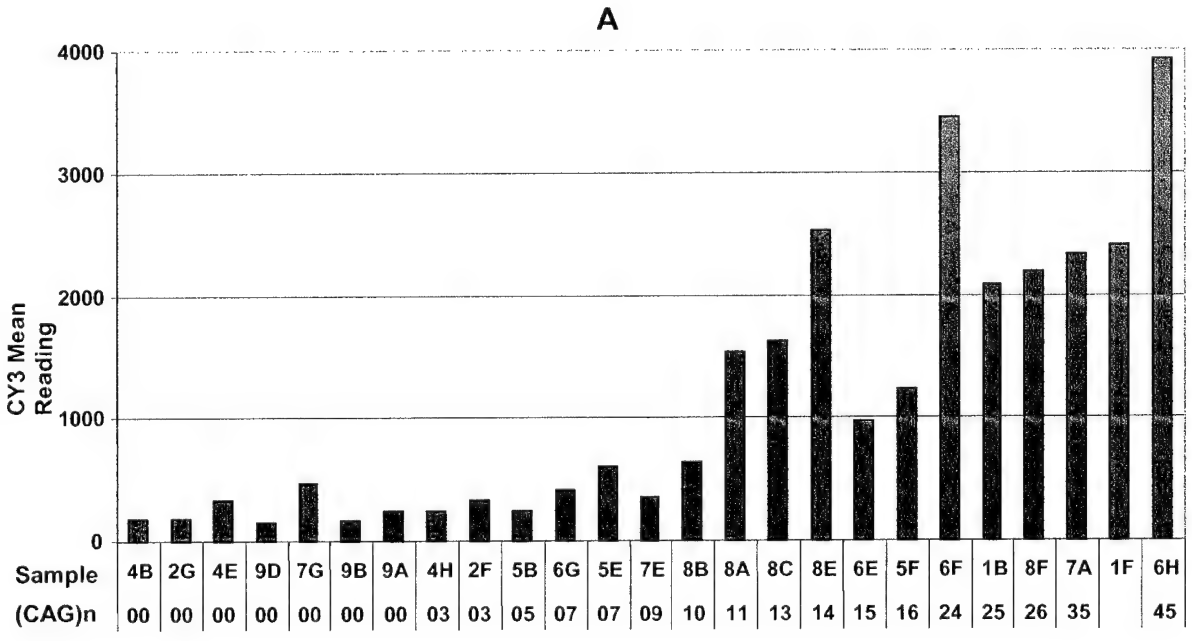


Figure 4: Correlation between CAG repeat length and signal intensity obtained by RepeatArray analysis. Both $(CAG)_7$ and $(CAG)_{17}$ probe provided similar results.



4. Bioinformatics Analyses of the Identified Repeats

In the initial screening 48 out of the 80 sequenced samples were found to contain (CAG) repeats ranging in size from $(CAG)_3$ to $(CAG)_{65}$. The sequences flanking the repeat region of the eight inserts which have (CAG) repeat larger than $(CAG)_{24}$ were blasted against the human genome sequence database in order to localize them and search for genes in the vicinity of repeat regions **(Table 2)**. We noticed that in some insert sequences the 5' flanking sequence of the repeat was blasted to a given chromosomal location and the 3' flanking sequence of the repeat was blasted to a different chromosomal location. This is due to the formation of chimeric repeat containing fragments during one of the PCR steps involved in the dynabead enrichment protocol due to sequence homology in the repeat regions. Recently we are working on overcoming this problem by increasing the stringency of the annealing step in the PCR protocols and adding some denaturing reagents to the PCR reaction mixtures.

Table 2: Genomic locations and nearby gene information of the repeat containing DNA fragments identified by Dynabead enrichment protocol

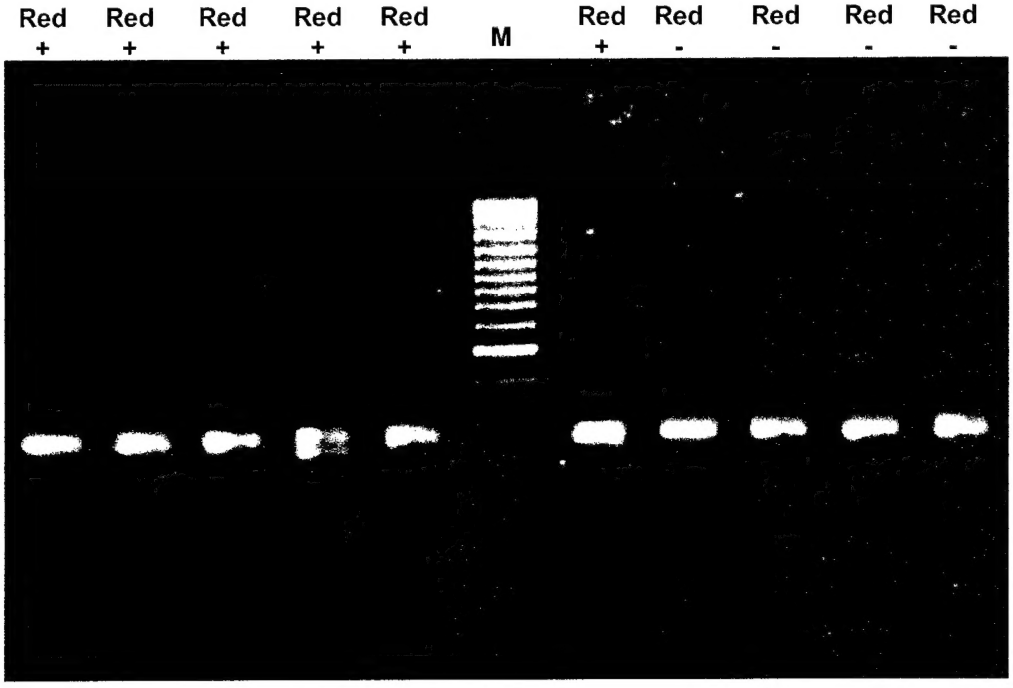
Clone	Length (bp)	(CAG/CTG) Length	Chromosome	Nearest gene
6F	850	24	Chr 14	unknown
7H	850	25	Chr 20	Chr20 NCOA6
1B	1300	25	Chr 3 Chr 5	Ch 3 ALCAM Ch 5 Model :hmm33519
8F	650	26	Chr 13 Chr 16	Chr 13 MAB21L1 (CAG)19 at 5'UTR of the gene
7A	600	35	Chr 22 Chr 12	Hs.440689
6H	750	45	Chr 2	DPP10,
1F	800	46	Chr 7	Foxp2 repeat is in exon
8H	700	65	Chr 2:	Unknown

5. Determining the Allelic Status of the large Repeat Regions

After identifying and localization of the repeat flanking sequences in the large (CAG) containing clones we designed primer pairs for both sides of the repeat of interest using Oligo software and we screened the panel of patient samples which had previously revealed to carry expanded CAG repeats by RED method. This strategy enables us to validate the size of the repeat between RED positive patient and control specimens. In an example presented in Figure 5, there were no significant differences in the length of PCR products between expanded and control samples **(Figure 5)**. This means that the expanded (CAG) repeats which were detected by RED are not due to the identified

sequences. Using this procedure we will continue to study the other candidate repeat regions suggested by the cloning method.

Figure 5. PCR products using the primers designed from the from $(CAG)_{65}$ containing inserts RED positive and RED negative samples



C. Key Research Accomplishments

At this stage, we have developed a reliable cloning strategy by integrating several enrichment steps and RepeatArray system for efficient identification of large repeat containing fragments. The specific technical accomplishments include:

(a) Establishment of the dynabead enrichment cloning protocol which increased the percentage of identifying repeat containing fragments to about 90% with an enrichment of larger repeats. The efficiency of the method has been validated.

(b) Establishment of the RepeatArray method for efficient screening for (CAG) repeats containing sequences. The method also provides the distinction between small and large repeats, which provides a high throughput pre-screening tool and eliminates the need for excess sequencing.

(c) Establishment of the protocol for determining the allelic status of the candidate repeats

(d) Preliminary characterization of the location of the candidate repeats relative to nearby gene sequences using Bioinformatics

D. Immediate Future Task

(a) Refine the dynabead enrichment protocol by increasing stringency of annealing and washing steps in order to eliminate the formation of chimeric inserts which are produced as an artifact in the PCR steps.
(b) Applying the recently optimized Dynabead enrichment and repeat array screening system to all RED positive patients samples in order to localize and characterize the expanded (CAG) repeats in those samples.
(c) Using bioinformatic tools, the functional characterization of the repeats will be carried out on the long (CAG) repeat containing loci in order to elucidate their biological significance and their influence on genes in their vicinity.
(d) As the sequences flanking the expanded repeats are identified, primers will be designed and the polymorphism status of the identified loci will be determined in 100 breast cancer and control samples.
(e) Search for information published on the loci/genes found and their association with cancer specifically with breast cancer.
(g) Writing of manuscripts to be proposed to peer-reviewed journals

E. Reportable Outcomes

Hamdi Jarjanazi, Hong Li, Irene L. Andrulis, and Hilmi Ozcelik, "Cloning strategy for trinucleotide repeat expansion: Searching for novel breast cancer predisposition gene(s)". Controversies in the etiology, detection and treatment of breast cancer: 2002, June 13-14, 2002 Toronto, Ontario, Canada *(Poster Presentation)*.

Hamdi Jarjanazi, Hong Li, Irene L. Andrulis, and Hilmi Ozcelik, "Identification of novel breast cancer predisposition gene(s) with trinucleotide repeat expansions". The Samuel Lunenfeld Research Institute Annual Retreat: 2002: October, 9-10 YMCA Geneva Park, Ontario, Canada *(Poster Presentation)*.

Hamdi Jarjanazi, Hong Li, Irene L. Andrulis, and Hilmi Ozcelik, "Identification of novel breast cancer predisposition gene(s) with trinucleotide repeat expansions". AACR 2003 Annual Meeting: 2002, April 5-9, Toronto, Ontario Canada *(Mini-symposium Presentation)*.

Hamdi Jarjanazi, Hong Li, Irene L. Andrulis and Hilmi Ozcelik, "Systematic Approach to Study the Role of Repeat Sequences in Cancer-Related Genes." (Poster), Proceedings of the American Association of Cancer Research (AACR) 95th Annual meeting, Orlando, FL, USA, March 2004 (*Poster Presentation*)

Hamdi Jarjanazi, Hong Li, Irene L. Andrulis, and Hilmi Ozcelik, " Genome Wide Screening of CAG/CTG Trinucleotide Repeat Lengths in Breast Cancer" *(manuscript in preparation)*

Hamdi Jarjanazi, Noel Pabalan, Hilmi Ozcelik. Rapid and efficient cloning of long (CAG) repeat sequences using dynabeads enrichment and DNA Array hybridization techniques (*manuscript in preparation*)

Hamdi Jarjanazi, Hilmi Ozcelik. Systematic approach to study the role of repeat sequences in cancer-related genes (*manuscript in preparation*).

References

Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. Biotechniques. 1999 Sep;27(3):500-2, 504-7

Inoue S, Takahashi K, Ohta M. Sequence analysis of genomic regions containing trinucleotide repeats isolated by a novel cloning method. Genomics. 1999 Apr 1;57(1):169-72.

Koob MD, Benzow KA, Bird TD, Day JW, Moseley ML, Ranum LP. Rapid cloning of expanded trinucleotide repeat sequences from genomic DNA. Nat Genet 1998 Jan;18(1):72-5

Margolis RL, McInnis MG, Rosenblatt A, Ross CA. Trinucleotide repeat expansion and neuropsychiatric disease. Arch Gen Psychiatry 56(11):1019-31, 1999.

O'Donovan MC, Guy C, Craddock N, Bowen T, McKeon P, Macedo A, Maier W, Wildenauer D, Aschauer HN, Sorbi S, Feldman E, Mynett-Johnson L, Claffey E, Nacmias B, Valente J, Dourado A, Grassi E, Lenzinger E, Heiden AM, Moorhead S, Harrison D, Williams J, McGuffin P, Owen MJ.Confirmation of association between expanded CAG/CTG repeats and both schizophrenia and bipolar disorder. Psychol Med 26(6):1145-53, 1996.

Pieretti M, Zhang FP, Fu YH, Warren ST, Oostra BA, Caskey CT and Nelson DL. Absence of expression of the FMR-1 gene in fragile X syndrome. Cell 66:817-822, 1991.

Sanjeeva P and Housman DE. The complex pathology of trinucleotide repeats. Current Opinion in Cell Biology 9:364-372, 1997.

Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H, Wakisaka A, Tashiro K, Ishida Y, Ikeuchi T, Koide R, Saito M, Sato A, Tanaka T, Hanyu S, Takiyama Y, Nishizawa M, Shimizu N, Nomura Y, Segawa M, Iwabuchi K, Eguchi I, Tanaka H, Takahashi H, Tsuji S. Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. Nat Genet. 1996 Nov;14(3):277-84.

Trottier Y, Lutz Y, Stevanin G, Imbert G, Devys D, Cancel G, Saudou F, Weber C, David G, Tora L, et al. Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. Nature 378:403-406, 1995.

Vincent JB, Neves-Pereira ML, Paterson AD, Yamamoto E, Parikh SV, Macciardi F, Gurling HM, Potkin SG, Pato CN, Macedo A, Kovacs M, Davies M, Lieberman JA, Meltzer HY, Petronis A, Kennedy JL., An unstable trinucleotide-repeat region on chromosome 13 implicated in spinocerebellar ataxia: a common expansion locus Am J Hum Genet 2000 Mar;66(3):819-29

Yuan QP, Lindblad-Toh K, Zander C, Burgess C, Durr A, Schalling M. A, cloning strategy for identification of genes containing trinucleotide repeat expansions. Int J Mol Med. 2001 Oct;8(4):427-31